Published online 11 September 2023

NAR Cancer, 2023, Vol. 5, No. 3 **1**
https://doi.org/10.1093/narcan/zcad047

Large-scale phenogenomic analysis of human cancers uncovers frequent alterations affecting SMC5/6 complex components in breast cancer

Shamayita Roy[1], Arvin Zaker[2], Arvind Mer[2,*] and Damien D'Amours[1,*]

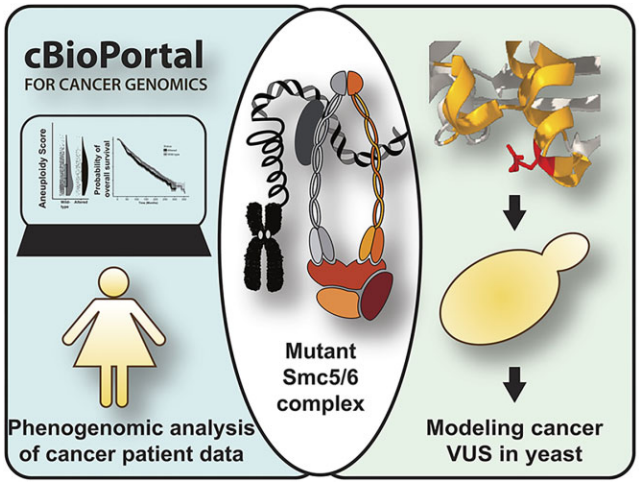
[1]Ottawa Institute of Systems Biology, Department of Cellular and Molecular Medicine, University of Ottawa, Roger Guindon Hall, 451 Smyth Rd, Ottawa, ON K1H 8M5, Canada and [2]Department of Biochemistry, Microbiology & Immunology, University of Ottawa, Roger Guindon Hall, 451 Smyth Rd, Ottawa, ON K1H 8M5, Canada

Received February 22, 2023; Revised August 09, 2023; Editorial Decision August 14, 2023; Accepted August 22, 2023

ABSTRACT

Cancer cells often experience large-scale alterations in genome architecture because of DNA damage and replication stress. Whether mutations in core regulators of chromosome structure can also lead to cancer-promoting loss in genome stability is not fully understood. To address this question, we conducted a systematic analysis of mutations affecting a global regulator of chromosome biology –the SMC5/6 complex– in cancer genomics cohorts. Analysis of 64 959 cancer samples spanning 144 tissue types and 199 different cancer genome studies revealed that the SMC5/6 complex is frequently altered in breast cancer patients. Patient-derived mutations targeting this complex associate with strong phenotypic outcomes such as loss of ploidy control and reduced overall survival. Remarkably, the phenotypic impact of several patient mutations can be observed in a heterozygous context, hence providing an explanation for a prominent role of SMC5/6 mutations in breast cancer pathogenesis. Overall, our findings suggest that genes encoding global effectors of chromosome architecture can act as key contributors to cancer development in humans.

GRAPHICAL ABSTRACT



INTRODUCTION

Accurate replication, repair and segregation of genomic DNA are essential processes for normal cell growth and maintenance of genome stability. Failure to execute these processes accurately leads to genome instability and disease states in humans and other organisms (1). In particular, a compelling body of evidence indicates that numerical as well as structural alterations of chromosomes can act as key initiating events in the development of cancer and other severe illnesses (1–4). Thus, proteins that play significant roles in upholding genome stability –namely, DNA repair and cell cycle checkpoint proteins– are often the ones that are found to be mutated in the genomes of cancer patients (5,6). It is perhaps surprising, however, that much less is known regarding the role of global effectors of chromosome structure/morphogenesis in cancer development, especially when considering the prevalence of

*To whom correspondence should be addressed. Tel: +1 613 562 5213; Email: Damien.Damours@uottawa.ca
Correspondence may also be addressed to Arvind Mer. Tel: +1 613 562 5213; Email: amer@uottawa.ca

chromosome-level changes in the genomes of cancer patients (e.g. ∼90% or more aneuploidy).

Structural Maintenance of Chromosome (SMC) protein complexes are responsible for the large-scale organization of the genetic material within cells (reviewed in (7–10)). The SMC family includes the well-characterized cohesin (SMC1/3) and condensin (SMC2/4) complexes as well as the more recently identified SMC5/6 complex. In eukaryotes, cohesin maintains proximity and alignment of sister chromatids whereas condensin is associated with the compaction/morphogenesis of chromosomes during cell division (7,10). Recent studies have shown that several cancer types harbor rare mutations in cohesin and condensin complexes supporting their role in suppression of tumorigenesis (11–13). Whether they contribute more broadly to oncogenesis remains to be established.

The SMC5/6 complex has important roles in diverse cellular processes affecting chromosome integrity throughout the cell cycle. Its main functions include the repair of DNA double strand break and stabilization of stalled replication forks (reviewed in (8,9)). Furthermore, the NSMCE2/Mms21 subunit of the SMC5/6 complex sumoylates multiple chromosome proteins via its E3 SUMO ligase activity (9,14–16) and is involved in the alternate lengthening of telomeres (ALT) maintenance pathway (17,18). We (and others) have found that the SMC5/6 core complex is a structure specific DNA binding enzyme (19–23) that promotes genome stability via its DNA compaction activity (24,25). Due to its major role in upholding genomic integrity, misregulation of the SMC5/6 complex often result in abnormal cellular growth and severe chromosomal abnormalities (reviewed in (8,9)). This raises the intriguing possibility that alterations in the components of the SMC5/6 complex might be associated with cancer development and other tumorigenic processes in humans. So far, very few reports have explored whether misregulation of the SMC5/6 complex might be directly responsible for cancer development in humans (17,26). The components of the SMC5/6 complex have been shown to be upregulated or altered in some cancer types (e.g. hepatocellular carcinoma (27,28), sarcoma (29), breast cancer (30) and brain metastasis (31)). However, it is currently unknown if these changes are passenger mutations or play a greater role in triggering oncogenesis in patients. This highlights the need for systematic studies to better define the significance of SMC5/6 complex alterations in cancer etiology and patient survival.

To fill this knowledge gap, we performed an integrative analysis of cancer genomic data (32) to identify the frequency and nature of cancer-associated alterations in the components of the SMC5/6 complex. Our phenogenomics analyses revealed that genes encoding the subunits of the SMC5/6 complex are frequently altered in breast as well as other types of human cancers. Importantly, these alterations are associated with severe ploidy aberrations and lower overall survival rates in patients. Extensive validation of the phenotypic effects associated with genetic variants of unknown significance (VUS) revealed a range of mutational severity, from mild to severe growth defects for point mutations affecting conserved functional domains, to lethality with alleles that truncate SMC5/6 components. Our results unravel a previously underappreciated pattern of frequent alterations in the components of the SMC5/6 complex in cancer genomes.

MATERIALS AND METHODS

Cancer cohort analysis

Cancer genomics and patient survival data was obtained from the cBioPortal for Cancer Genomics (https://cbioportal.org) web resources on August 2022 (32). Obtained data contained 64 959 samples from 199 studies and 145 tissue types (577 tissue subtypes). During the quality control steps, duplicated samples, pediatric cancer cohorts and cohorts with samples size <20 were removed. Next, we selected patient samples for which mutation, copy-number alteration (CNA), and structural variation (SV) information was available. The final dataset contains 29 316 samples from 33 tissue types and 100 studies. Data processing, analysis and visualization was performed on R statistical software (version 4.2.2) using packages ggplot2 (version 3.4.0), ComplexHeatmap (version 2.14.0) and circlize (version 0.4.15). In the analysis, the SMC5/6 complex was considered altered if any of the genes encoding its constituents (*NSMCE1*, *NSMCE2*, *NSMCE3*, *NSMCE4A*, *SMC5*, *SMC6* and *EID3*) had alteration. Survival analysis was performed using R package survival (version 3.4-0) and survminer (version 0.4.9). Kaplan–Meier plots were used for visualization of survival differences and log-rank-test was used to estimate statistical significance.

Variation allele frequency analysis of patients with SMC5/6 complex mutations was performed on the TCGA dataset. Mutation allele frequency files were downloaded using the TCGAbiolinks (version 2.25.3) package (33) for the samples containing point-mutation in SMC5/6 complex components. We then calculated the variation allele frequency (VAF) using the formula:

$$\mathrm{VAF} = \frac{f_{\mathrm{altered}}}{f_{\mathrm{altered}} + f_{\mathrm{reference}}}$$

where f_{altered} is the frequency of the DNA sequence reads with the genomic alteration, and $f_{\mathrm{reference}}$ is the frequency of the reference DNA sequencing reads without genomic alterations. A variant is generally considered to be homozygous if its VAF is >50%, meaning that more than half of the sequencing reads support the variant allele. We then visualized the VAF values of all patients with recorded SMC5/6 complex alteration using ggplot2 (version 3.4.0).

Differential gene expression and pathway analysis

Differential gene expression analysis of microarray data was performed using limma package (version 3.52.4) (34). For RNAseq based gene expression data, edgeR package (version 3.40.0) (35) was used, and data were fitted using the GLM method (35). Gene expression data for METABRIC (36) study was obtained using the MetaGxBreast package (version 1.18.0) (37). STAR aligner (38) based The Cancer Genome Atlas (TCGA) (39) RNA-Seq data was downloaded using the TCGAbiolinks package (version 2.25.3) (33).

Gene ranking based on the log fold change values were utilized for gene set enrichment analysis. We obtained the

canonical pathways (version 7.5.1) from the Molecular Signatures Database (MSigDB). Pathways with number of genes <25 and >150 were removed from the analysis. R package fgsea (version 1.25.0) was used for gene set enrichment analysis and statistical significance (*P*-value) for each pathway was estimated using adaptive multi-level split Monte-Carlo scheme (40). False discovery rate (FDR) was used for *P*-value adjustment of multiple hypothesis testing. Pathways were categorized into groups such as DNA replication and repair, cancer-specific, cell cycle-related, metabolic, signaling, extracellular matrix, immune pathway based on semantic analysis (41). Enriched pathway network was constructed where node represents pathways and edge represents number of shared genes between pathways. Nodes were connected only when more than 20 genes were shared between nodes. Visualization of network was done using igraph package (version 1.3.5).

Comparative analysis of SMC5/6 alterations with known effectors of genomic instability

Co-mutational pattern of SMC5/6 complex was analyzed against several genomic stability related genes. Mutation, copy-number alteration (CNA), and structural variation (SV) information for these genes were downloaded from the cBioPortal web server. We then compared the co-mutational patterns of the SMC5/6 complex against the genomic instability genes using the co-mutation frequency table. For the analysis we computed false discovery rate (FDR) (42) of co-mutation pattern. FDR was defined as:

$$\mathrm{FDR} = \frac{\mathrm{FP}}{\mathrm{TP} + \mathrm{FP}}$$

where FP is false positive (number of samples with alteration in genomic instability gene and wild-type SMC5/6 gene), and TP is true positive (number of samples with alteration in both genomic instability and SMC5/6 genes). This analysis was performed for each SMC5/6 complex gene and genomic stability gene pair. Pie chart was used for co-mutation frequency visualization.

We also compared the transcriptional signature associated with SMC5/6 complex alteration against genomic instability related transcriptional signatures. For this analysis, we first performed differential gene-expression analysis between samples with high and low genomic instability. If at least three of 12 essential genomic stability genes were mutated, cancer samples were labeled as exhibiting high genomic instability. Log fold change (LFC) values associated with genomic stability were compared to LFC values associated with SMC5/6 complex changes for all genes.

Mutation profile analysis

Lollipop plots of different mutation profiles for respective genes in all cancers were generated by the mutation mapper in cBioPortal. Evolutionary conservation of the mutations in SMC5/6 complex subunits were compared in 5 species, *Saccharomyces cerevisiae* (budding yeast), *Saccharomyces pombe* (fission yeast), *Xenopus laevis* (frog), *Drosophila melanogaster* (fruit fly), *Homo sapiens* (human). The amino-acid sequence of eukaryotic homologs of SMC5/6 complex subunits were acquired from Uniprot (https://www.uniprot.org/). The amino acid sequences of the genes were aligned in the Clustal Omega sequence alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). Additionally, the conserved amino-acid residues were shaded with Boxshade online tool (https://embnet.vitalit.ch/software/BOX_form.html).

Yeast strains and cell viability assay

All yeast strains used in this study are derivatives of strain K699/K700. The genotype of the yeast strains used in the study are listed in supplementary information (Supplementary Table S1). Yeast growth conditions, media composition and procedures for genetic analysis can be found elsewhere (43). For experiments performed under conditions of DNA damage or replication stress, yeast cultures were grown on solid medium containing methyl methanesulfonate (MMS), 4-nitroquinoline 1-oxide (4-NQO) and hydroxyurea (HU) at 23°C, 30°C and 37°C. Temperature and DNA damage sensitivity were monitored for all mutants created in this study because several DNA repair/homologous recombination mutants show sensitivity to both types of stress (44,45). In summary, 5-fold dilution series of wild-type and mutant yeast cultures (first spot on the left side of the plate corresponds to a culture at OD_{600} of 0.2) were spotted on solid YPD (yeast extract, peptone, 2% glucose) and grown in temperature-controlled incubators for 48–72 h before scanning the plates in a scanner (46). All the experiments were repeated a minimum of three independent times, and we show representative results in figures.

Plasmid and mutant construction

Yeast strains expressing cancer-specific gene truncations were created by integration of a T_{ADH1}-*kanMX6 or* T_{ADH1}*::URA3MX6* cassette at the desired point of truncation in the endogenous loci of genes encoding the SMC5/6 complex subunits. Heterozygous diploid strains carrying these truncation alleles were then sporulated on minimal media and the viability of the dissected haploid spores was determined after 4 days of growth on solid medium. Missense mutations were initially introduced in plasmids (*pFA6a-ORF::*T_{ADH1}*::kanMX6* or *pFA6a-ORF::*T_{ADH1}*::URA3MX6*) carrying *SMC5*, *SMC6*, *NSE1*, *MMS21*, *NSE3* or *NSE4* genes using QuikChange Multi Mutagenesis kit (Agilent). Mutant alleles of the respective genes were then transformed into wildtype diploid yeast strain at the relevant endogenous loci by transformation of a PCR product from the region of interest in the respective mutagenized plasmids. Heterozygous diploid strains carrying these missense mutation alleles were then sporulated and dissected to determine the phenotype of the mutant haploid strains. All the mutations in the haploid mutants were confirmed by sequencing the respective genetic loci.

Immunoblot analysis

Cell lysates were prepared from exponential cultures of yeast grown at 23°C or 37°C using the TCA glass-bead

method (47). Lysates were subsequently resolved by SDS-PAGE and processed for immunoblot analysis using anti-Mcm4 antibody (clone C-12; 1:500 dilution; Santa Cruz Biotechnology; sc-166036), anti-Myc antibody (clone 9E10; 1:2500 dilution; Cedarlane; GTX20032), anti-Pgk1 antibody (clone 22C5D8; 1:2500 dilution; Abcam) and an anti-mouse IgG antibody (1:5000 dilution; Cytiva). Band intensity on immunoblots was measured using Adobe Photoshop (version 2022). Data are presented as means ± SEM. All statistical analyzes were performed using GraphPad Prism 7 (GraphPad Software Inc.) and statistical significance threshold was set at P-value = 0.05. Where indicated in figure legends, we performed multiple t-tests comparing the mutants to the wild-type for either temperature, P-values were corrected by Bonferroni multiple correction method.

RESULTS

Subunits of SMC5/6 complex are frequently altered across different cancer types

The components of the human SMC5/6 complex assemble into a ring-like complex similar in configuration to that of cohesin and condensin, the defining members of this family of proteins (8,9) (Figure 1A). The ring structure of the complex is formed by the dimerization of two SMC proteins, SMC5 and SMC6, and their association with a group of accessory subunits named NSMCE1, NSMCE3 and NSMCE4A/EID3. Additionally, the human SMC5/6 complex comprises an E3 SUMO ligase, NSMCE2 (also known as Mms21/Nse2) (14,15), which interacts with the coiled-coil domain of SMC5 (Figure 1A) (48,49).

We analyzed 64 959 cancer samples spanning across 144 tumor types and 199 different cancer genome studies for evidence of genetic alterations in the seven genes encoding the known components of the SMC5/6 complex (Figure 1A). Genomic alterations including structural variation (SV), mutations, and copy-number alterations (CNA), at the SMC5/6 complex were investigated. After data preprocessing and quality control, we derived 29 316 high-quality cancer sample profiles across 34 tumor types including lung cancer ($N = 6745$), breast cancer ($N = 6709$), and prostate cancer ($N = 2791$) (Figure 1B and Supplementary Figure S1). Analysis of the curated dataset revealed that genomic alteration of the SMC5/6 complex is present in 10.9% of all cancer samples.

Stratifying by tissue of origin revealed that genetic alterations affecting SMC5/6 complex components are most prevalent in ovarian cancer (21.4%) followed by breast (18.1%) and endometrial cancer (14.5%). Analysis of cooccurrence patterns revealed that the majority of patients have alteration in only one gene encoding members of the SMC5/6 complex and *NSMCE2*/*NSMCE1* have high frequency of coexisting mutation (Figures 1C and 2). The general trend of SMC5/6 subunit alterations we observed in individual cancers (Figure 2) is similar to the trend we observed in aggregate (Figure 1). Namely, amplification of *NSMCE2* is the most common genomic alteration, except in endometrial cancer where missense mutations are prevalent (Supplementary Figure S1). Genomic alteration of *NSMCE2* is dominant in breast, prostate, and ovarian cancer while alteration in *SMC5* and *SMC6* genes are highly prevalent in lung, melanoma, bladder, and endometrial cancer (Figure 2). To determine the zygosity of SMC5/6 complex mutations, we conducted a variant allele frequency (VAF) analysis of cancer alterations affecting the subunits of the complex. Only 0.4% of samples had biallelic mutations (VAF > 0.5) in any of the SMC5/6 complex subunits, showing that the vast majority of SMC5/6 complex mutations are heterozygotes (Figure 3A; one tail t-test P-value < 0.0001).

Next, we evaluated whether mutations in the SMC5/6 complex are directly linked to mutations in genomic stability-related genes (Figure 3B). We analyzed co-mutations patterns between genes encoding the SMC5/6 complex and key genes associated with genome stability and computed false discovery rate (FDR) for each combination. We did not observe statistically significant co-mutation patterns (all FDR > 0.05). We discovered that the *TP53-SMC5/6* complex co-mutation rate was the highest. This was expected (and not statistically significant) given the high prevalence of *TP53* mutation in cancer. These results strongly suggest that genetic alterations affecting the SMC5/6 complex are not directly driven by mutations in genomic stability-related genes.

Genomic alteration in SMC5/6 complex is associated with aggressive disease

To understand the effect of SMC5/6 complex alterations on aggressive cancer, we performed an in-depth analysis of its association with ploidy score and survival outcome (Figure 4). We focused on breast cancer, as it has the highest number of samples with the SMC5/6 complex alterations (Figure 2). Analysis of breast cancer genomes revealed that individuals harboring the SMC5/6 complex alteration have significantly higher aneuploidy scores (Figure 4A, Wilcoxon signed-rank test P-value $= 3 \times 10^{-4}$) and ploidy scores (Figure 4B, Wilcoxon signed-rank test P-value $= 1 \times 10^{-6}$). This association with high aneuploidy and ploidy scores was also observed in cancer genomes from individuals carrying *NSMCE2* alterations. Analysis of male-specific prostate cancer data showed a similar trend (Supplementary Figure S2), suggesting that high aneuploidy and ploidy scores are likely due to SMC5/6 complex alterations. Survival analysis using Kaplan–Meier (KM) plots shows that patients with altered SMC5/6 complex had significantly lower overall survival rates compared to patients with normal SMC5/6 complex components (Figure 4C, Log-rank P-value $= 6.0 \times 10^{-4}$). We also observed that the *NSMCE2* alteration is associated with significantly worse patient survival relative to individuals carrying wild-type *NSMCE2* (Figure 4D, Log-rank P value $= 1.0 \times 10^{-5}$).

We observed a consistent pattern while stratifying the genomic alterations into categories. The amplification of SMC5/6 complex genes were associated with higher aneuploidy and ploidy score and poor patient survival (Supplementary Figure S3). The same holds true for NSMCE2 gene amplification. Point mutation to ploidy score analysis was not possible due to a lack of data. However, survival analysis unequivocally demonstrates that mutations

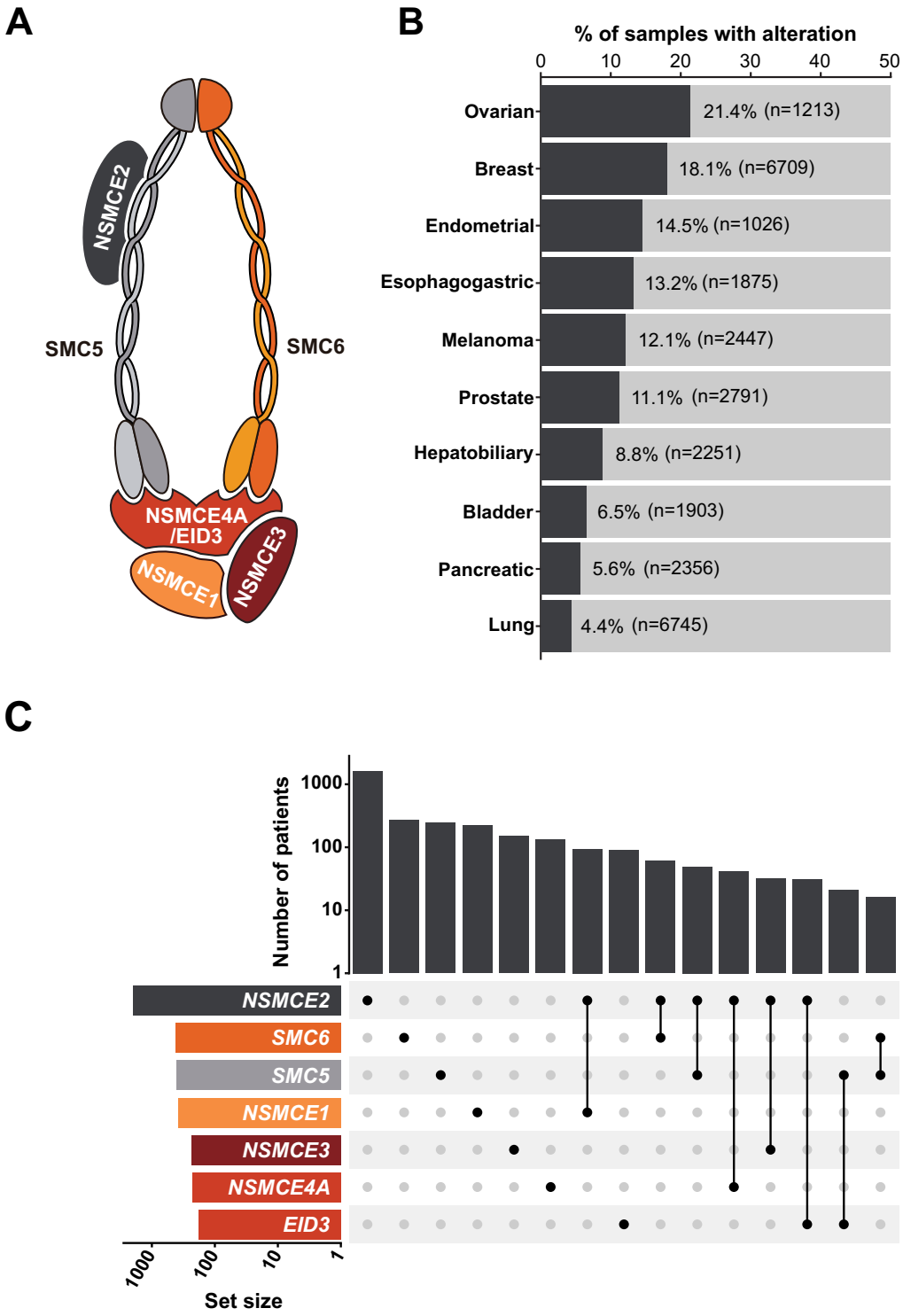


Figure 1. Genetic alterations affecting the subunits of the SMC5/6 complex in human cancers. (**A**) Schematic representation of the SMC5/6 complex. Note that the NSMCE4A subunit is replaced by EID3 in male gonads (88). (**B**) Bar chart showing the frequency of SMC5/6 complex alteration in cancer. For visualization, 10 tissue types with the highest frequency of SMC5/6 complex alteration were chosen. (**C**) UpSet plot of the distribution of genomic alteration of each subunit of the SMC5/6 complex in the 10 major cancer tissue types reported herein. Number of patients (y-axis) are shown using logarithmic scale.

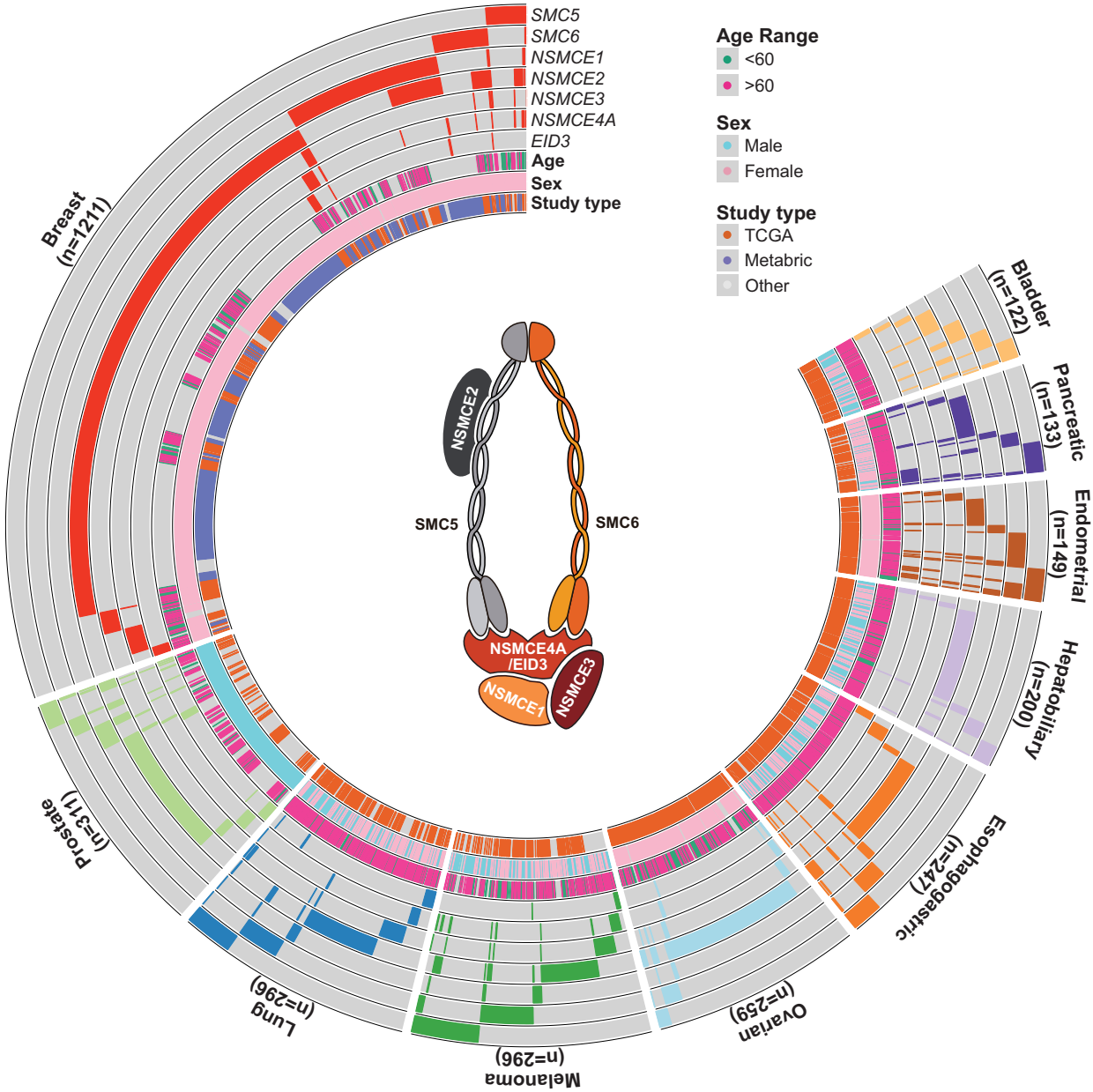


Figure 2. Circular oncoprint of patients carrying genomic alteration in the SMC5/6 complex. Each patient is represented by a cross-section of the circle. Recorded genomic alteration in each of the subunits is shown as a color relating to the cancer type. Age is dichotomized at 60 years. The number of patients with genomic alteration in SMC5/6 complex belonging to each cancer type is shown under the name.

in the SMC5/6 complex result in worse prognosis, even when compared to amplification (Supplementary Figure S3B). The difference in overall patient survival in cohorts of patients carrying wild-type and altered SMC5/6 complex components was also observed in individuals with prostate and ovarian cancer (Supplementary Figures S2 and S4, respectively). Overall, our analyses show that the alterations in the genes encoding the SMC5/6 complex are linked to high cellular ploidy and are associated with poor prognosis with *NSMCE2* being the most frequent target of alteration.

SMC5/6 complex alteration are linked to DNA damage and replication stress

Differential gene expression analysis was performed to discover genes that are linked with SMC5/6 complex alteration in breast cancer (Figure 5A and Supplementary Figure S5). We found that several key genes involved in DNA damage repair, cell cycle and DNA replication were differentially expressed in the SMC5/6 complex alteration group (Figure 5B). Gene *RAD54B*, a member of DEAD-like helicase superfamily was upregulated in SMC5/6 complex alteration

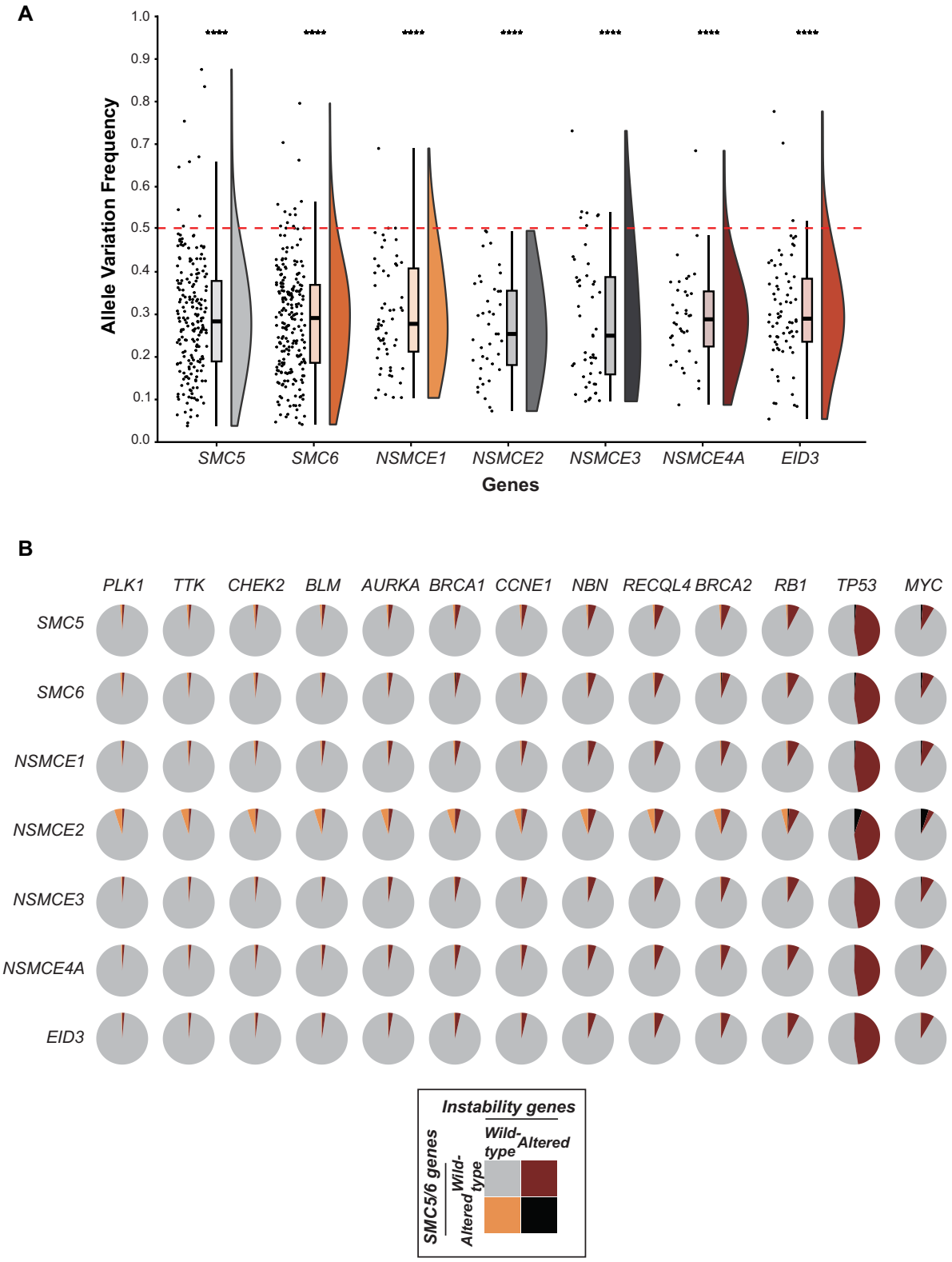


Figure 3. (**A**) Zygosity of SMC5/6 complex alterations in human cancers. Different SMC5/6 complex genes are shown on the X-axis of the graph, whereas the Y-axis shows variant allele frequency (VAF) in cancer samples. Most mutations have VAF < 0.5 indicating that they are heterozygotes. Statistical significance was computed using one tail *t*-test (H1: average VAF is < 0.5), and *P*-value < 0.0001, represented by **** at the top of each gene. (**B**) Co-mutation patterns between SMC5/6 complex and genes associated with genomic instability. Each row represents a component gene of the SMC5/6 complex, while columns represent gene related genomic instability. The mutational patterns are displayed as a pie chart, with the percentage of samples that have a wild-type status in both genes represented in the color gray, and the frequency of samples that have a mutation in both genes shown in the color black. The color orange indicates a mutation that has only occurred in the SMC5/6 complex, whereas the color dark red indicates a mutation that has only occurred in the genomic instability gene.

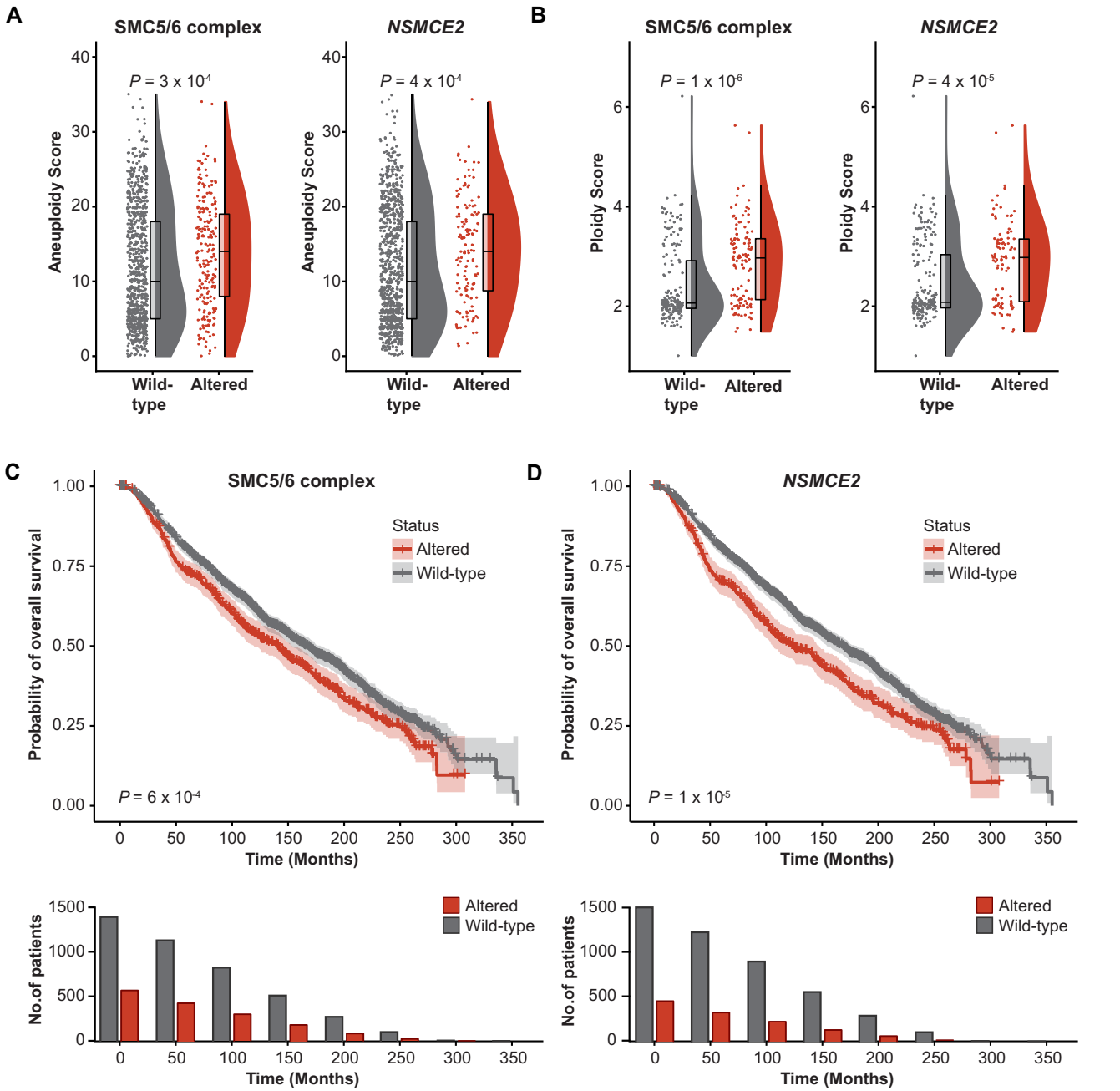


Figure 4. SMC5/6 complex alterations are associated with genome stability defects and reduced survival in breast cancer patients. Distribution of (**A**) aneuploidy and (**B**) ploidy score in breast cancer patients. Data are stratified by SMC5/6 complex and *NSMCE2* alteration status. Samples carrying SMC5/6 complex alterations have higher aneuploidy and ploidy score (two-sided Wilcoxon rank-sum test $P < 0.05$). Kaplan–Meier (KM) plots of overall survival in breast cancer, stratified by (**C**) SMC5/6 complex and (**D**) *NSMCE2* gene alteration status. Altered samples are shown in red and wild type in gray. The bar charts at the bottom of the KM plots represent the number of patients at risk. Alteration in SMC5/6 complex leads to poor overall survival (Log-rank test P-value < 0.05).

group (Figure 5B). It is known to play an active role in DNA damage repair and homologous recombination during cell division (50–52). *CCNE2* (Figure 5B), a cyclin family gene and vital component of G1/S transition during the cell cycle, showed high expression in SMC5/6 complex altered samples (53,54). Similarly, *RECQL4*, and *MCM4* (Figure 5B) were upregulated in SMC5/6 complex alteration group. These genes are associated with DNA repair (55) and replication (56) respectively. Genes such as *KIF13B*, *PPP2R2A*, *PARP3,* and *TP53BP1* were all significantly downregulated in SMC5/6 complex alteration group (Figure 5B). Interestingly, these genes have been implicated in the negative control of cell growth (57), maintenance of genomic stability (32) and tumor suppressor activity (58). Therefore, downregulation of these genes may lead to aggressive tumor phenotype.

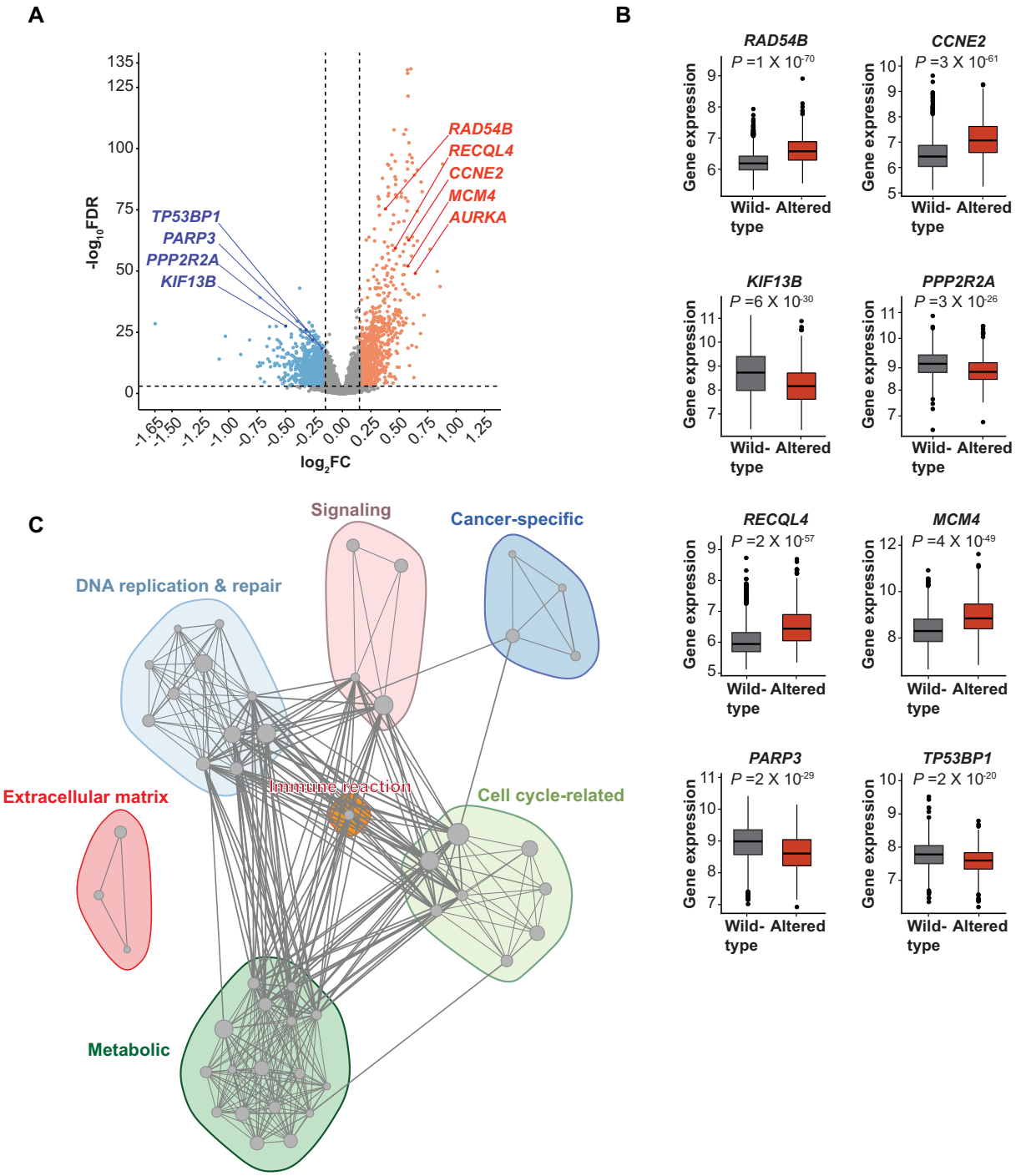


Figure 5. Genomic alteration in the SMC5/6 complex leads to changes in gene expression and pathway activity. (**A**) Volcano plot of the differential gene expression analysis of breast cancer data. In the volcano plot x-axis represents log fold change (LFC) and y-axis represents false discovery rate (FDR) for genes. Significantly upregulated genes are highlighted with orange and significantly downregulated genes are highlighted with blue. (**B**) Box plot showing the expression of selected genes in SMC5/6 complex altered (red) and wild-type (gray) breast cancer samples. Differential gene expression *P*-value is shown on the top. (**C**) Network plot of enriched pathways in breast cancers carrying altered SMC5/6 complex components. Each pathway is represented by a node. The size of the node is proportional to pathway size. Edges between two nodes indicate the number of genes shared between pathways. Only links with more than 20 shared genes are shown. Pathways are categorized into groups using semantic analysis and network layout is done using Fruchterman-Reingold algorithm. Detailed pathway information is available in Supplementary Figure S6.

We also compared the transcriptional signature of SMC5/6 complex modification to the transcriptional signature associated with genomic stability (Supplementary Figure S5A). We found that only 84 genes were differentially expressed in both the conditions (i.e. SMC5/6 complex alteration and mutation in key genomic stability genes). This constitutes only 3.72% of all genes that were differentially expressed in SMC5/6 complex. Consequently, 96.28% of the differentially expressed genes in SMC5/6 complex are unique to it. Furthermore, key genes such as *RAD54B*, *RECQL4,* and *MCM4* were uniquely linked to SMC5/6 complex alteration. Overall, these results established that the transcriptional changes associated with SMC5/6 complex alteration are independent of the genomic architecture abbreviations.

To explore the pathways associated with SMC5/6 complex alteration, we utilized gene-set enrichment analysis (GSEA) approach. Results revealed that the DNA replication and repair, cell cycle, metabolism, signaling and cancer related pathways are strongly associated with SMC5/6 complex alteration (Figure 5C and Supplementary Figure S6). A similar trend was observed in TCGA prostate cancer data (Supplementary Figure S5C). Taken together, these gene and pathway analyses results highlight a strong link between SMC5/6 complex alteration and DNA damage-related processes.

A large number of cancer mutations affect key functional domains of the SMC5/6 complex

In depth analysis of patient cancer genomes included in the study revealed that they harbor a large number of point mutations in the subunits of the SMC5/6 complex (Figure 6A). An analysis of the spatial distribution of mutations in the subunits of the SMC5/6 complex identified a total of 343 and 326 mutation sites in SMC5 and SMC6 proteins, respectively, with no major hotspot region in either of the proteins. R972* nonsense mutation was the most frequently observed alteration in the SMC5 subunit in multiple cancers. Also, the I410Yfs frameshift deletion in SMC6 was found to be ubiquitously present in different cancer types (Figure 6A). Although the non-SMC subunits of SMC5/6 complex were also mutated at several sites in different cancer types, no single mutation were recurring at a higher frequency compared to the SMC core subunits. Importantly, several of the mutations identified in the SMC5/6 core complex and the non-SMC subunits were present on amino-acid residues which are evolutionarily conserved across several eukaryotic species (Figure 6B). Furthermore, in many cancers missense mutations affect domains and/or structural modules of known functional relevance for the SMC5/6 complex (e.g. H187Y in the SP-RING domain of NSMCE2 and R229G/Q in the WH-B domain of NSMCE3) (Figure 6B). Although most, if not all point mutations reported here should be viewed as variants of unknown significance (VUS), it stands to reason that VUS affecting evolutionarily conserved regions of the SMC5/6 complex are likely to disrupt the activity of the complex. Since we do not know which exact biochemical function of the SMC5/6 complex is affected by the cancer mutations described above, it is important to investigate the functional impact of each mutation on SMC5/6 complex activity.

Modeling cancer VUS reveals the impact of cancer mutations on SMC5/6 complex activity

We next sought to assess the functional relevance of cancer VUS on the activity of the SMC5/6 complex. We took advantage of the high evolutionary conservation of the SMC5/6 complex in eukaryotes to model the impact of point mutations in the budding yeast *Saccharomyces cerevisiae*. To achieve this, we selected a group of cancer VUS affecting conserved amino acid residues at several positions within the subunits of the SMC5/6 complex, focusing on the known functional domains of each subunit (Figure 6). We then constructed a series of heterozygous diploid yeast strains carrying cancer VUS at their endogenous loci and uncovered the phenotype associated with the mutations of interest after sporulation and dissection of haploid spores (Figure 7 and Supplementary Table S2).

The majority of the haploid yeast strains carrying truncating VUS in the components of the SMC5/6 complex were non-viable after sporulation, as evidenced by the 2:2 lethality phenotype co-segregating with each truncation allele on dissection plates (Figure 7A). The only viable truncation mutant obtained in this analysis, *mms21-C221**, showed severe proliferation defects at 23°C and 30°C (relative to a wild-type control strain) and was completely defective for growth at 37°C (Figure 7B). Even at its permissive temperature of 23°C, the yeast strain carrying *mms21-C221** was unable to grow effectively on media containing DNA damaging agents (methyl methanesulfonate [MMS] and 4-nitroquinoline 1-oxide [4NQO]) or a DNA replication inhibitor (hydroxyurea [HU]; Figure 6B). The inability of the *mms21-C221** mutant strain to withstand DNA lesions as well as its temperature-sensitive growth defect are highly similar to that of a known DNA repair defective mutant of the complex, the *smc5-6* mutant (Figure 7B) (59).

Interestingly, all haploid yeast strains carrying missense mutations were viable after sporulation (Figure 7C) but showed variable kinetics of proliferation under standard and DNA damage conditions. While all the mutants grew normally on solid medium (YPD) at 23°C, *mms21-H202Y* and *nse1-C321R* mutants appeared to be temperature sensitive at 37°C and showed poor proliferation on media containing MMS or HU (Figure 7C). Likewise, the *smc6-G1021R* mutant was defective for growth at all temperatures tested when exposed to HU or MMS. Although yeast strains carrying the *smc5-E1015K*, *nse3-K236G* and *nse4-P315S* alleles showed normal growth phenotypes at the permissive temperature of 23°C, their proliferation was severely hindered when challenged with HU at the temperature of 37°C (Figure 7C). In summary, most of the mutant alleles we tested manifested defects in the functional activity of the SMC5/6 complex that resulted in mild to severely impaired growth phenotypes *in vivo.* A summarizing model of cancer point mutations and their impact on the functionality of the SMC5/6 complex is shown in Supplementary Figure S7.

Since, we observed that several key genes involved in DNA damage repair, cell cycle and DNA

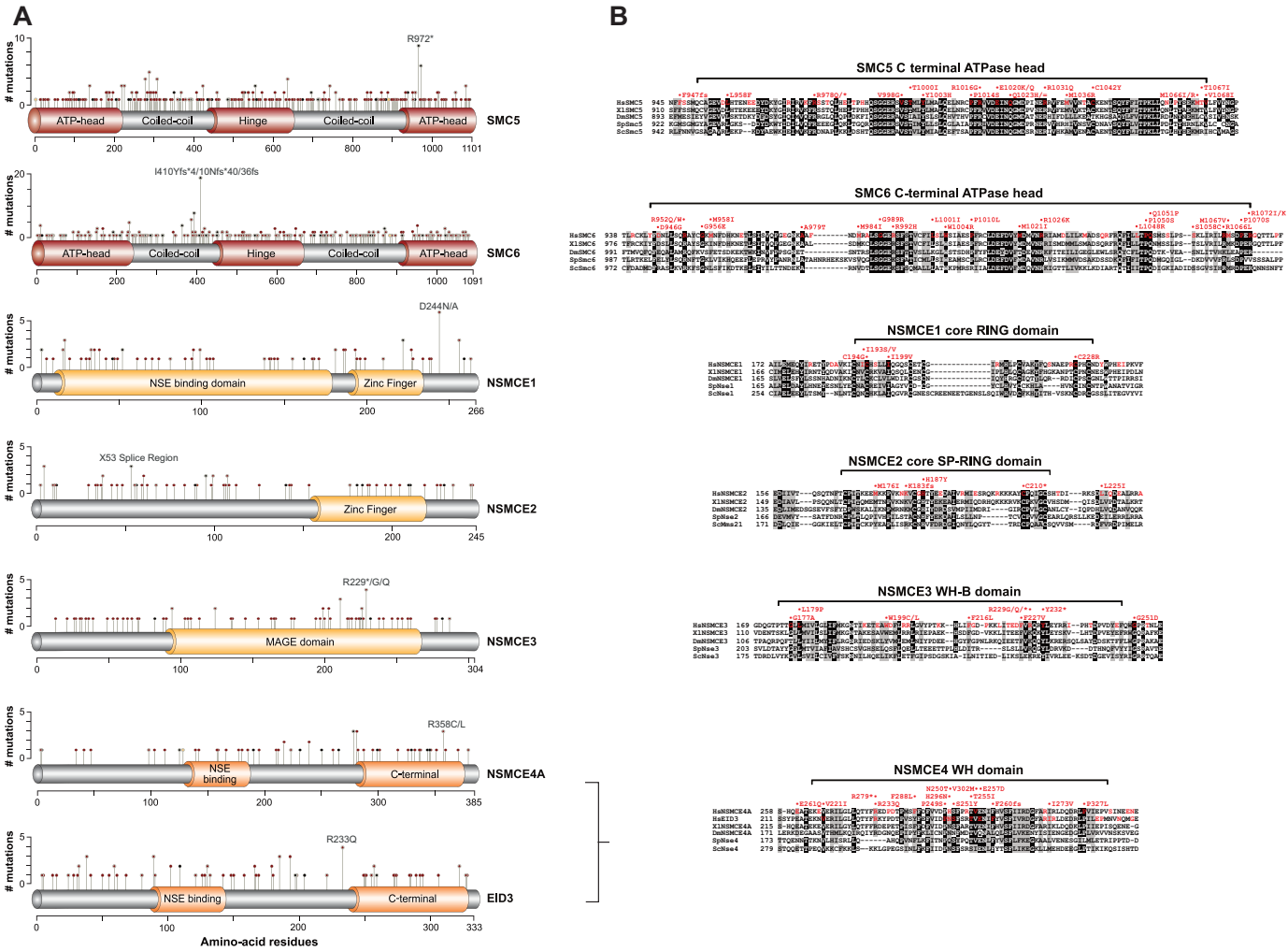


Figure 6. Cancer mutations in the genes encoding components of the SMC5/6 complex. (**A**) The relative positions of cancer mutations are marked on schematics depicting the domain organization of the subunits of the SMC5/6 complex. Specific protein names are marked on the right-hand side of the schematics. The length of the line connecting the annotated mutation to the protein is indicative of the number of samples that carry the mutation (frequency). For each subunit, we also report the specific nature of the mutations most frequently identified in cancers. Specifically: Black circles mark the positions for truncation mutations, red circles are for missense mutations and yellow circles are for frameshift mutations. (**B**) Position and evolutionary conservation of cancer mutations identified in functional domain of the SMC5/6 complex. Fully conserved residues are highlighted in black whereas positions conserved at $\geq$ 80% are highlighted in red. Mutations affecting positions that are evolutionarily conserved are marked in bold/red typeface, while mutations that affect non-conserved residues are marked in red typeface. Specific details pertaining to cancer mutations are marked above protein alignments. Mutational data are from cBioPortal. SMC5 C-terminal ATPase head region, SMC6 C-terminal ATPase head regions, NSMCE1 core RING domain, NSMCE2 core SP-RING domain, NSMCE3 winged helix B domain, and NSMCE4 winged helix domain.

replication—especially RAD54B and MCM4—are distinctively linked to SMC5/6 complex alteration and differentially expressed in the SMC5/6 complex alteration group (Figure 5B and Supplementary Figure S5), we wanted to test if the protein abundance of yeast Rdh54 (human RAD54B homolog) and yeast Mcm4 would be affected in strains carrying cancer-like mutations in the SMC5/6 complex (Supplementary Figure S8). Remarkably, mutation in glycine 1021 of Smc6 led to a detectable decrease in protein abundance of Mcm4 especially at non-permissive temperature of 37°C (Supplementary Figure S8A). Moreover, yeast strain carrying the *nse3-K236G* alleles had slight decrease in Rdh54 protein expression at 23°C (Supplementary Figure S8B). Together, these results indicate that cancer-like mutations in the yeast SMC5/6 complex reduce its ability to promote an effective response to DNA damage, a phenotype typically associated with reduced genome stability in eukaryotes.

Cancer mutations in SMC5/6 components are haploinsufficient for resistance to replication stress

The majority of the point mutations we identified in SMC5/6 complex components are present in only one allele of the corresponding gene pair found in cancer samples (Figure 3A). If they are to have a significant effect on cancer development, one would expect they can act in a dominant manner *in vivo* or, alternatively, that loss of a single allele of a given SMC5/6 component results in haploinsufficient phenotypes (60). To investigate these possibilities, we assessed the viability and growth properties of diploid yeast strains carrying wild-type and cancer VUS mutations

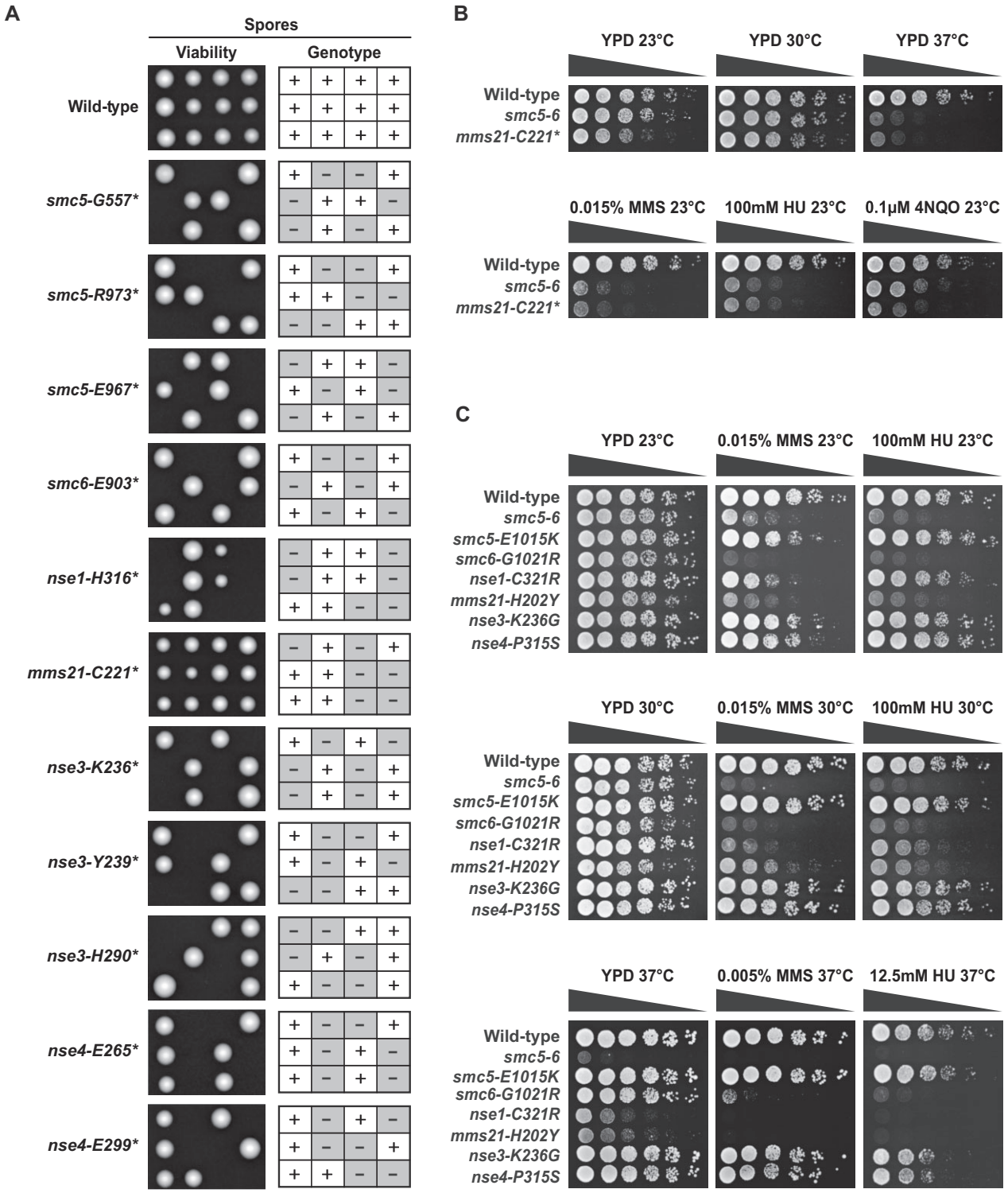


Figure 7. Modeling the phenotype of SMC5/6 cancer mutations in *Saccharomyces cerevisiae*. (**A**) The physiological impact of truncating mutations in SMC5/6 complex components was analyzed by sporulation and dissection of heterozygous diploid yeasts carrying specific mutations in SMC5/6 complex subunits. (**B**) Impact of the SUMO ligase truncation mutation –*mms21-C221**– in the yeast homolog of human NSMCE2. The proliferation capacity of yeast strains was monitored by dilution assay on solid medium and growth under combinations of genotoxic and high temperature conditions. (**C**) Functional impact of point mutations in the subunits of the SMC5/6 complex. Mutant strains were analyzed as in panel (B).

in a heterozygous context. Most of the yeast strains co-expressing wild-type and VUS alleles of SMC5/6 components showed marked growth defects when exposed to various forms of stress (Figure 8). For instance, all nonsense mutant alleles we tested in heterozygous diploid conditions showed significant growth defects at 37°C. These defects were substantially exacerbated in the presence of replication stress (Figure 8, middle and bottom HU panels). Interestingly, the defective growth behavior of strains carrying heterozygous truncation mutations was similar to that of strains carrying either a full deletion (*smc5Δ/SMC5*) or a previously identified conditional allele (*smc5-6/SMC5*) of *SMC5*. These results strongly suggest that expression levels of SMC5/6 complex components must be tightly regulated *in vivo* to allow full functionality of the complex, and that loss of 50% protein is sufficient to induce haploinsufficient growth defects in heterozygous strains. Consistent with this, we observed similar growth defects in cells expressing cancer missense mutations in SMC5/6 complex components (Figure 8, Top panel). These results dovetail nicely with the morphological haploinsufficiency phenotype observed in strains partly defective in the SMC5/6 complex (61). Taken together, these observations suggest that inactivation of a single allele of SMC5/6 complex components in cancer patients is sufficient to impair the activity of the complex and result in sensitivity to replicative stress.

We note that a cancer mutation identified in the *NSMCE1* gene behaved as a dominant mutation when introduced in yeast. Indeed, a heterozygous diploid strain expressing both the wild-type and *nse1-H316** truncation allele (i.e. corresponding to human *nsmce1-S222**) showed a dominant temperature-sensitive growth defect at 37°C (Figure 8, middle panel). The H316 position falls within an important Zinc finger domain in the Nse1 protein that can affect important functional activities like DNA binding, protein–protein interactions, and various other cellular functions (62). Interestingly, the corresponding amino acid residue in humans (Ser222 of NSMCE1) was found to be mutated in three individual cancer patients. This indicates that a subset of cancer mutations in SMC5/6 complex components may act in a dominant manner to inactivate the complex and promote tumorigenesis.

DISCUSSION

Loss of genome stability is a pivotal event that turns healthy cells towards cancerous growth. Whereas the cancer-promoting roles of mutations affecting DNA repair and checkpoint pathways are well established (5,6), the impact of chromosome-level defects on tumor development is much less understood. In particular, how defects in chromosome morphogenesis (i.e. individualization, compaction and morphology) during mitosis might drive oncogenic processes remain to be fully elucidated. The difficulty in deciphering the contribution of chromosome-level defects in cancer development stems in part from the fact that loss of global chromosome regulation during mitosis is typically associated with cell death, thus impeding detailed functional analysis (7–10). Moreover, when partial loss of global chromosome regulation is observed in human diseases—such as in individuals suffering from hypomorphic mutations in SMC complexes—the associated pathological consequences often lead to lethality at an age that precedes typical timelines of cancer development (63–65). These limitations have considerably hampered progress in understanding how global defects in chromosome morphogenesis might impact cancer development in humans.

To address the challenge outlined above, we took advantage of mutation-rich datasets provided by large-scale efforts to sequence human cancer genomes (32) to understand how alterations in the genes encoding global chromosome regulators might impact oncogenesis. We focused our analysis on the SMC5/6 complex since this is one of the least well understood effector of chromosome architecture in eukaryotes (8,9). Remarkably, our analysis revealed that components of the SMC5/6 complex are frequently altered in several types of human cancers, including breast (18.05%) and prostate cancers (11.14%). Our data also shows that cancers carrying co-mutations in multiple subunits of SMC5/6 complex are rare (1.6%) and alterations in single SMC5/6 complex genes are not associated with patient age and sex.

Remarkably, our results establish that alterations in genes encoding the components of the SMC5/6 complex lead to loss of genome stability (i.e. higher ploidy score) and adverse patient survival outcome while demonstrating no detectable co-mutational patterns with a panel of genes frequently connected to genome instability. These results are consistent with a recent report by Grange *et al.* (65) showing that a pathogenic variant of the SMC5/6 complex associates with hyperploidy during embryonic development. Our analysis suggests that alterations in the SMC5/6 complex causes hyperploidy in both breast and prostate cancers. DNA ploidy is known to be prognostic to patient survival and a high ploidy score is an indicator of worse prognosis (66,67). Collectively, these results indicate that the alteration in SMC5/6 complex causes global changes in the stability of cancer genomes, leading to an aggressive disease phenotype and poor patient prognosis.

Our findings also indicate that one of the most frequently affected components of the SMC5/6 complex in human cancer is NSMCE2, a SP-RING domain protein with E3 SUMO ligase activity (14,15). Previous experiments have shown that this activity is important for DNA repair reactions (14,15,68) and our modeling experiments in budding yeast show that cancer point mutations in the SP-RING domain of Mms21 interfere with the maintenance of genome stability after DNA damage. Interestingly, it has been shown that removal of *NSMCE2* in human osteosarcoma (U2OS cells) and human breast cancer (ER^+/PR^+: MCF-7 cells) resulted in significant differences in phenotypes, including slower cell growth and cell cycle arrest (69,70). Additionally, increased telomerase activity associated with heightened cell survival through activation of ALT pathways, is detected in 90% of prostate carcinomas (71) and can stem from the amplification of *NSMCE2* gene. In alignment with this finding, we also observe ubiquitous amplification of *NSMCE2* in a large fraction of prostate and breast cancers reported in our study. It is noteworthy that the genomic locus for *NSMCE2* is located near the *MYC* locus on chromosome 8. We have noticed that *NSMCE2* gene is often co-amplified along with the *MYC* gene in an amplicon containing region 8q24 in cancer pa-

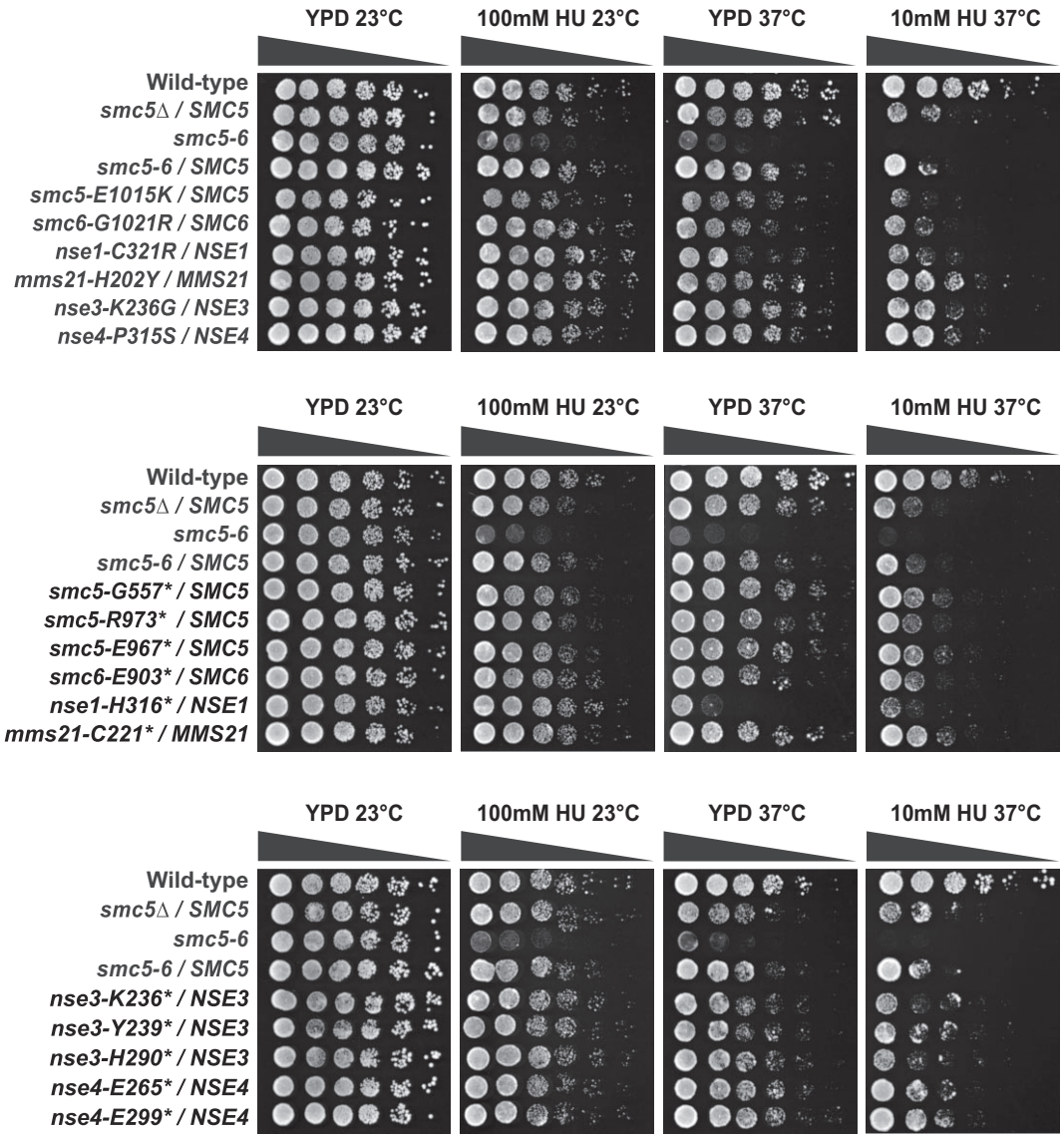


Figure 8. Haploinsufficent phenotype of SMC5/6 cancer mutations in budding yeast. The impact of specific cancer mutations in the subunits of the SMC5/6 complex was assessed in heterozygous diploid yeasts. The proliferation capacity of yeast strains was monitored by dilution assay on solid medium in the presence of genotoxic agents and/or temperature stress, as described in the legend.

tients. While it is clear that an increase in *MYC* abundance can act as a driver for cancer formation, it's amplification on double minutes (dmin) does not always lead to *MYC* overexpression (72) suggesting that other genes in the 8q24 amplicon can contribute to oncogenesis. When *NSMCE2* is coamplified with *MYC* in the same cell, we envision that this would create a highly potent 'double hit' scenario because of the synergy associated with simultaneous stimulation of MYC's oncogenic properties and loss of genome stability (due to SMC5/6 complex inactivation). Our results provide primary evidence for this hypothesis. In the MYC-amplified group, NSMCE2 alteration was associated with poor patient survival, demonstrating that simultaneous NSMCE2 and MYC modification is associated with a bleak prognosis (Supplementary Figure S3A, log-rank P-value = 0.03). This result further demonstrates the influence of NSMCE2 alterations on patient survival, independent of MYC alterations. Considering the important role of genome instability as a hallmark of cancer (73), it seems highly probable that misregulation of the SMC5/6 complex can promote cancer development as either an independent driver of the process or a cooperator with other oncogenes (such as *MYC*).

Comprehensive analysis of genomic data from cancer patient databases revealed that point mutations and partial deletions occur frequently in all the subunits of the SMC5/6 complex. Prior to our analysis, the functional and tumorigenic impact of these genetic alterations was unknown, hence the 'variant of unknown significance' (VUS) status of SMC5/6 mutations in cancer databases. To shed some light on their functional impact and potential role in cancer development, we introduced several of these cancer-specific mutations at their homologous positions in the subunits of SMC5/6 complex in yeast. Remarkably, all but one

SMC5/6 subunit truncations led to lethality in yeast after sporulation and dissection of heterozygous diploid strains. Only the deletion of the C-terminus of yeast Mms21 (human NSMCE2 homolog) led to a viable but sick yeast strain suggesting that even a modest loss of amino-acid sequence can result in a strong impairment of SMC5/6 complex function (14).

Our genetic analysis of diploid yeast strains carrying both wild-type and cancer-specific mutations support the view that expression of a single mutant allele of the SMC5/6 complex is sufficient to weaken the DNA damage response of yeast. Consistent with this, reducing the abundance of Smc5 by half led to a detectable proliferation defect in the *smc5Δ/SMC5* heterozygous mutant, suggesting that mild imbalances in the expression of SMC5/6 complex components can have an impact on the functionality of the complex (74). Similar results have been observed in plants overexpressing *HPY2* (*Arabidopsis thaliana MMS21/NSMCE2*) (75,76). It is noteworthy that misregulation (including overexpression) of *HPY2* directly impacts ploidy control in plants (hence the name '*High Ploidy2*' for mutants of the *HPY2* gene), an important aspect of genome stability in all eukaryotes (76). This observation dovetails nicely with the hyperploidy phenotype reported by Grange and colleagues in Atelís syndrome patients carrying mutations in the components of the SMC5/6 complex (65). Overall, our data suggest that dual allele inactivation of SMC5/6 complex components is not necessary to promote genome instability and tumorigenesis in patients. Haploinsufficient mutations in genes regulating the DNA damage response, such as *ATM* and *BLM*, are well known to drive cancer formation (77,78), which provides a compelling paradigm to directly implicate single-allele mutations of the SMC5/6 complex in cancer development. Consistent with this view, haplo-insufficiency of *NSMCE2* was recently shown to be associated with higher incidence of tumor formation and overall poor survival in a murine model of cancer (26). Taken together, this data supports the view that clinically relevant single-allele mutations affecting evolutionary conserved domains of the SMC5/6 complex or overexpression of some of its components likely contribute as driver mutations for cancer initiation and/or progression.

How might defects in the SMC5/6 complex contribute to cancer initiation and/or maintenance? Hints of the possible answer to this question come from our observation that SMC5/6-defective cancers show dramatically altered ploidy levels. It is now widely appreciated that changes in the chromosome contents of cells, in particular aneuploidy (1–4), create a substantial burden on many cellular processes and can promote cancer development. For instance, high level of aneuploidy is a major driving force for prostate cancer development, is indicative of a higher extent of aggressiveness in primary prostate cancers (79) and can confer resistance to chemotherapy (80,81). Our work revealed that patients with altered SMC5/6, particularly with altered *NSMCE2*, have high aneuploidy scores in prostate and breast cancers, which outlines a likely mechanism by which SMC5/6 complex mutations can promote cancer development. Interestingly, defects in SMC5/6 activity might also provide a unique window of opportunity for cancer treatment. Extensive analysis of genetic interaction networks in yeast have identified several synthetic lethal interactions connecting members of the SMC5/6 complex with proteins involved in the resolution of DNA replication damage. Many of these proteins have clear homologs in higher eukaryotes and are known to be mutated in several human cancers (26,82,83). Since many genetic interactions are conserved in eukaryotes, we propose that synthetic lethal partners of the SMC5/6 complex –particularly those involved in the repair of DNA replication damage– are likely to be effective targets for therapeutic intervention in cancers harboring mutations in the SMC5/6 complex. This prediction is further supported by our observation that yeast strains carrying cancer-specific mutations in its SMC5/6 complex are highly sensitive to replication-stress induced by HU. Since HU and its derivatives have been shown to be effective anti-cancer drugs with minimal side effects (84–87), we predict cancers harboring mutations in SMC5/6 complex components would respond well to this drug or other types of DNA replication stressors (e.g. CHK1 inhibitors) (87).

In conclusion, our work demonstrates the pervasiveness of mutations affecting the SMC5/6 complex in diverse human cancers and strongly suggests that patients harboring SMC5/6-defective cancers would benefit from precision/genotype-targeted treatment modalities focused on creating synthetic lethality connected to DNA replication stress. Future efforts will be required to faithfully manipulate specific functional pathways connected with the SMC5/6 complex and leverage that knowledge to develop better diagnostic procedures and therapeutic agent to treat human cancers.

DATA AVAILABILITY

Patient data including mutation, DNA ploidy score and survival information used in this study are available at the cBioPortal for Cancer Genomics (https://cbioportal.org) web resources. METABRIC (36) study gene expression data is available through MetaGxBreast package (version 1.18.0) (37). Processed next generation sequencing based gene expression data is available through TCGAbiolinks package (version 2.25.3) (41). All necessary code to reproduce the bioinformatic analyses and figures described in this manuscript are available on the Zenodo open repository site at https://doi.org/10.5281/zenodo.8256896. All accompanying data can be found on Zenodo at https://doi.org/10.5281/zenodo.8256823.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Cancer Online.

ACKNOWLEDGEMENTS

We thank members of the D'Amours laboratory for their comments on the manuscript, and Dr Hemanta Adhikary in particular for his help with the structural model of the SMC5/6 complex.

Author contributions: S.R. and D.D. conceived and designed yeast experiments; A.Z. and A.M. performed all bioinformatics analyses; S.R. created yeast strains, assessed their viability and DNA damage sensitivity; S.R. and A.Z. prepared figures; A.Z. and A.M. wrote all the bioinformatics

sections of the paper; S.R. and D.D. wrote the sections of the paper pertaining to yeast experiments.

FUNDING

CIHR [FDN-167265 to D.D.]; J.P. Bickell Foundation Medical Research Grant (to A.M.); D.D. is supported by a Canada Research Chair in Chromatin Dynamics & Genome Architecture.
Conflict of interest statement. None declared.

REFERENCES

1. Keuper,K., Wieland,A., Räschle,M. and Storchova,Z. (2021) Processes shaping cancer genomes - from mitotic defects to chromosomal rearrangements. *DNA Repair (Amst.)*, **107**, 103207.
2. Kneissig,M., Bernhard,S. and Storchova,Z. (2019) Modelling chromosome structural and copy number changes to understand cancer genomes. *Curr. Opin. Genet. Dev.*, **54**, 25–32.
3. Garribba,L. and Santaguida,S. (2022) The dynamic instability of the aneuploid genome. *Front. Cell Dev. Biol.*, **10**, 838928.
4. Ben-David,U. and Amon,A. (2020) Context is everything: aneuploidy in cancer. *Nat. Rev. Genet.*, **21**, 44–62.
5. Groelly,F.J., Fawkes,M., Dagg,R.A., Blackford,A.N. and Tarsounas,M. (2023) Targeting DNA damage response pathways in cancer. *Nat. Rev. Cancer*, **23**, 78–94.
6. Klinakis,A., Karagiannis,D. and Rampias,T. (2020) Targeting DNA repair in cancer: current state and novel approaches. *Cell. Mol. Life Sci.*, **77**, 677–703.
7. Yatskevich,S., Rhodes,J. and Nasmyth,K. (2019) Organization of chromosomal DNA by SMC complexes. *Annu. Rev. Genet.*, **53**, 445–482.
8. Aragón,L. (2018) The Smc5/6 complex: new and old functions of the enigmatic long-distance relative. *Annu. Rev. Genet.*, **52**, 89–107.
9. Solé-Soler,R. and Torres-Rosell,J. (2020) Smc5/6, an atypical SMC complex with two RING-type subunits. *Biochem. Soc. Trans.*, **48**, 2159–2171.
10. Kim,Y. and Yu,H. (2020) Shaping of the 3D genome by the ATPase machine cohesin. *Exp. Mol. Med.*, **52**, 1891–1897.
11. De Koninck,M. and Losada,A. (2016) Cohesin mutations in cancer. *Cold Spring Harb. Perspect. Med.*, **6**, a026476.
12. Yuen,K.C. and Gerton,J.L. (2018) Taking cohesin and condensin in context. *PLoS Genet.*, **14**, e1007118.
13. Leiserson,M.D.M., Vandin,F., Wu,H.-T., Dobson,J.R., Eldridge,J.V., Thomas,J.L., Papoutsaki,A., Kim,Y., Niu,B., McLellan,M. *et al.* (2015) Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.*, **47**, 106–114.
14. Zhao,X. and Blobel,G. (2005) A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 4777–4782.
15. Potts,P.R. and Yu,H. (2005) Human MMS21/NSE2 is a SUMO ligase required for DNA repair. *Mol. Cell Biol.*, **25**, 7021–7032.
16. Dhingra,N. and Zhao,X. (2021) Advances in SUMO-based regulation of homologous recombination. *Curr. Opin. Genet. Dev.*, **71**, 114–119.
17. Potts,P.R. and Yu,H. (2007) The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nat. Struct. Mol. Biol.*, **14**, 581.
18. Noël,J.-F. and Wellinger,R.J. (2011) Abrupt telomere losses and reduced end-resection can explain accelerated senescence of Smc5/6 mutants lacking telomerase. *DNA Repair (Amst.)*, **10**, 271–282.
19. Taschner,M., Basquin,J., Steigenberger,B., Schäfer,I.B., Soh,Y.-M., Basquin,C., Lorentzen,E., Räschle,M., Scheltema,R.A. and Gruber,S. (2021) Nse5/6 inhibits the Smc5/6 ATPase and modulates DNA substrate binding. *EMBO J.*, **40**, e107807.
20. Yu,Y., Li,S., Ser,Z., Sanyal,T., Choi,K., Wan,B., Kuang,H., Sali,A., Kentsis,A., Patel,D.J. *et al.* (2021) Integrative analysis reveals unique structural and functional features of the Smc5/6 complex. *Proc. Natl. Acad. Sci. U.S.A.*, **118**, e2026844118.
21. Yu,Y., Li,S., Ser,Z., Kuang,H., Than,T., Guan,D., Zhao,X. and Patel,D.J. (2022) Cryo-EM structure of DNA-bound Smc5/6 reveals DNA clamping enabled by multi-subunit conformational changes. *Proc. Natl. Acad. Sci. U.S.A.*, **119**, e2202799119.
22. Hallett,S.T., Schellenberger,P., Zhou,L., Beuron,F., Morris,E., Murray,J.M. and Oliver,A.W. (2021) Nse5/6 is a negative regulator of the ATPase activity of the Smc5/6 complex. *Nucleic Acids Res.*, **49**, 4534–4549.
23. Adamus,M., Lelkes,E., Potesil,D., Ganji,S.R., Kolesar,P., Zabrady,K., Zdrahal,Z. and Palecek,J.J. (2020) Molecular insights into the architecture of the Human SMC5/6 complex. *J. Mol. Biol.*, **432**, 3820–3837.
24. Serrano,D., Cordero,G., Kawamura,R., Sverzhinsky,A., Sarker,M., Roy,S., Malo,C., Pascal,J.M., Marko,J.F. and D'Amours,D. (2020) The Smc5/6 core complex is a structure-specific DNA binding and compacting machine. *Mol. Cell*, **80**, 1025–1038.
25. Gutierrez-Escribano,P., Hormeño,S., Madariaga-Marcos,J., Solé-Soler,R., O'Reilly,F.J., Morris,K., Aicart-Ramos,C., Aramayo,R., Montoya,A., Kramer,H. *et al.* (2020) Purified Smc5/6 complex exhibits DNA substrate recognition and compaction. *Mol. Cell*, **80**, 1039–1054.
26. Jacome,A., Gutierrez-Martinez,P., Schiavoni,F., Tenaglia,E., Martinez,P., Rodríguez-Acebes,S., Lecona,E., Murga,M., Méndez,J., Blasco,M.A. *et al.* (2015) NSMCE2 suppresses cancer and aging in mice independently of its SUMO ligase activity. *EMBO J.*, **34**, 2604–2619.
27. Nie,H., Wang,Y., Yang,X., Liao,Z., He,X., Zhou,J. and Ou,C. (2021) Clinical significance and integrative analysis of the SMC Family in hepatocellular carcinoma. *Front. Med. (Lausanne)*, **8**, 727965.
28. Yang,H., Gao,S., Chen,J. and Lou,W. (2020) UBE2I promotes metastasis and correlates with poor prognosis in hepatocellular carcinoma. *Cancer Cell Int.*, **20**, 234.
29. Zhou,J., Wu,G., Tong,Z., Sun,J., Su,J., Cao,Z., Luo,Y. and Wang,W. (2020) Prognostic relevance of SMC family gene expression in human sarcoma. *Aging (Albany NY)*, **13**, 1473–1487.
30. Di Benedetto,C., Oh,J., Choudhery,Z., Shi,W., Valdes,G. and Betancur,P. (2022) NSMCE2, a novel super-enhancer-regulated gene, is linked to poor prognosis and therapy resistance in breast cancer. *BMC Cancer*, **22**, 1056.
31. Saunus,J.M., Quinn,M.C.J., Patch,A.-M., Pearson,J.V., Bailey,P.J., Nones,K., McCart Reed,A.E., Miller,D., Wilson,P.J., Al-Ejeh,F. *et al.* (2015) Integrated genomic and transcriptomic analysis of human brain metastases identifies alterations of potential clinical significance. *J. Pathol.*, **237**, 363–378.
32. Gao,J., Aksoy,B.A., Dogrusoz,U., Dresdner,G., Gross,B., Sumer,S.O., Sun,Y., Jacobsen,A., Sinha,R., Larsson,E. *et al.* (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal*, **6**, pl1.
33. Colaprico,A., Silva,T.C., Olsen,C., Garofano,L., Cava,C., Garolini,D., Sabedot,T.S., Malta,T.M., Pagnotta,S.M., Castiglioni,I. *et al.* (2016) TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.*, **44**, e71.
34. Ritchie,M.E., Phipson,B., Wu,D., Hu,Y., Law,C.W., Shi,W. and Smyth,G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **43**, e47.
35. Chen,Y., Lun,A.T.L. and Smyth,G.K. (2016) From reads to genes to pathways: differential expression analysis of RNA-seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res*, **5**, 1438.
36. Curtis,C., Shah,S.P., Chin,S.-F., Turashvili,G., Rueda,O.M., Dunning,M.J., Speed,D., Lynch,A.G., Samarajiwa,S., Yuan,Y. *et al.* (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, **486**, 346–352.
37. Gendoo,D.M.A., Zon,M., Sandhu,V., Manem,V.S.K., Ratanasirigulchai,N., Chen,G.M., Waldron,L. and Haibe-Kains,B. (2019) MetaGxData: clinically annotated breast, ovarian and pancreatic cancer datasets and their use in generating a multi-cancer gene signature. *Sci. Rep.*, **9**, 8770.
38. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
39. Cancer Genome Atlas Research Network, Weinstein,J.N., Collisson,E.A., Mills,G.B., Shaw,K.R.M., Ozenberger,B.A., Ellrott,K., Shmulevich,I., Sander,C. and Stuart,J.M. (2013) The cancer genome atlas pan-cancer analysis project. *Nat. Genet.*, **45**, 1113–1120.

40. Korotkevich,G., Sukhov,V., Budin,N., Shpak,B., Artyomov,M.N. and Sergushichev,A. (2021) Fast gene set enrichment analysis. bioRxiv doi: https://doi.org/10.1101/060012, 01 February 2021, preprint: not peer reviewed.
41. Yu,G., Wang,L.-G., Yan,G.-R. and He,Q.-Y. (2015) DOSE: an R/bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*, **31**, 608–609.
42. Broer,L., Lill,C.M., Schuur,M., Amin,N., Roehr,J.T., Bertram,L., Ioannidis,J.P.A. and van Duijn,C.M. (2013) Distinguishing true from false positives in genomic studies: p values. *Eur. J. Epidemiol.*, **28**, 131–138.
43. St-Pierre,J., Douziech,M., Bazile,F., Pascariu,M., Bonneil,E., Sauvé,V., Ratsima,H. and D'Amours,D. (2009) Polo kinase regulates mitotic chromosome condensation by hyperactivation of condensin DNA supercoiling activity. *Mol. Cell*, **34**, 416–426.
44. Sinha,H., David,L., Pascon,R.C., Clauder-Münster,S., Krishnakumar,S., Nguyen,M., Shi,G., Dean,J., Davis,R.W., Oefner,P.J. *et al.* (2008) Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. *Genetics*, **180**, 1661–1670.
45. Auesukaree,C., Damnernsawad,A., Kruatrachue,M., Pokethitiyook,P., Boonchird,C., Kaneko,Y. and Harashima,S. (2009) Genome-wide identification of genes involved in tolerance to various environmental stresses in Saccharomyces cerevisiae. *J. Appl. Genet.*, **50**, 301–310.
46. Ratsima,H., Ladouceur,A.-M., Pascariu,M., Sauvé,V., Salloum,Z., Maddox,P.S. and D'Amours,D. (2011) Independent modulation of the kinase and polo-box activities of Cdc5 protein unravels unique roles in the maintenance of genome stability. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, E914–E923.
47. Foiani,M., Marini,F., Gamba,D., Lucchini,G. and Plevani,P. (1994) The B subunit of the DNA polymerase alpha-primase complex in Saccharomyces cerevisiae executes an essential function at the initial stage of DNA replication. *Mol. Cell. Biol.*, **14**, 923–933.
48. Duan,X., Sarangi,P., Liu,X., Rangi,G.K., Zhao,X. and Ye,H. (2009) Structural and functional insights into the roles of the Mms21 subunit of the Smc5/6 complex. *Mol. Cell*, **35**, 657–668.
49. Varejão,N., Lascorz,J., Codina-Fabra,J., Bellí,G., Borràs-Gas,H., Torres-Rosell,J. and Reverter,D. (2021) Structural basis for the E3 ligase activity enhancement of yeast Nse2 by SUMO-interacting motifs. *Nat. Commun.*, **12**, 7013.
50. Xu,C., Liu,D., Mei,H., Hu,J. and Luo,M. (2019) Knockdown of RAD54B expression reduces cell proliferation and induces apoptosis in lung cancer cells. *J. Int. Med. Res.*, **47**, 5650–5659.
51. Miyagawa,K., Tsuruga,T., Kinomura,A., Usui,K., Katsura,M., Tashiro,S., Mishima,H. and Tanaka,K. (2002) A role for RAD54B in homologous recombination in human cells. *EMBO J.*, **21**, 175–180.
52. Wesoly,J., Agarwal,S., Sigurdsson,S., Bussen,W., Van Komen,S., Qin,J., van Steeg,H., van Benthem,J., Wassenaar,E., Baarends,W.M. *et al.* (2006) Differential contributions of mammalian Rad54 paralogs to recombination, DNA damage repair, and meiosis. *Mol. Cell. Biol.*, **26**, 976–989.
53. Payton,M., Scully,S., Chung,G. and Coats,S. (2002) Deregulation of cyclin E2 expression and associated kinase activity in primary breast tumors. *Oncogene*, **21**, 8529–8534.
54. Caldon,C.E., Sergio,C.M., Sutherland,R.L. and Musgrove,E.A. (2013) Differences in degradation lead to asynchronous expression of cyclin E1 and cyclin E2 in cancer cells. *Cell Cycle*, **12**, 596–605.
55. Lu,H., Shamanna,R.A., Keijzers,G., Anand,R., Rasmussen,L.J., Cejka,P., Croteau,D.L. and Bohr,V.A. (2016) RECQL4 Promotes DNA end resection in repair of DNA double-strand breaks. *Cell Rep.*, **16**, 161–173.
56. Ishimi,Y., Komamura-Kohno,Y., Kwon,H.-J., Yamada,K. and Nakanishi,M. (2003) Identification of MCM4 as a target of the DNA replication block checkpoint system. *J. Biol. Chem.*, **278**, 24644–24650.
57. Kalev,P., Simicek,M., Vazquez,I., Munck,S., Chen,L., Soin,T., Danda,N., Chen,W. and Sablina,A. (2012) Loss of PPP2R2A inhibits homologous recombination DNA repair and predicts tumor sensitivity to PARP inhibition. *Cancer Res.*, **72**, 6414–6424.
58. Squatrito,M., Vanoli,F., Schultz,N., Jasin,M. and Holland,E.C. (2012) 53BP1 is a haploinsufficient tumor suppressor and protects cells from radiation response in glioma. *Cancer Res.*, **72**, 5250–5260.
59. Torres-Rosell,J., Machín,F., Farmer,S., Jarmuz,A., Eydmann,T., Dalgaard,J.Z. and Aragón,L. (2005) SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. *Nat. Cell Biol.*, **7**, 412–419.
60. Veitia,R.A. (2002) Exploring the etiology of haploinsufficiency. *Bioessays*, **24**, 175–184.
61. Ohnuki,S. and Ohya,Y. (2018) High-dimensional single-cell phenotyping reveals extensive haploinsufficiency. *PLoS Biol.*, **16**, e2005130.
62. Pebernard,S., Perry,J.J.P., Tainer,J.A. and Boddy,M.N. (2008) Nse1 RING-like domain supports functions of the Smc5-Smc6 holocomplex in genome stability. *Mol. Biol. Cell*, **19**, 4099–4109.
63. van der Crabben,S.N., Hennus,M.P., McGregor,G.A., Ritter,D.I., Nagamani,S.C.S., Wells,O.S., Harakalova,M., Chinn,I.K., Alt,A., Vondrova,L. *et al.* (2016) Destabilized SMC5/6 complex leads to chromosome breakage syndrome with severe lung disease. *J. Clin. Invest.*, **126**, 2881–2892.
64. Payne,F., Colnaghi,R., Rocha,N., Seth,A., Harris,J., Carpenter,G., Bottomley,W.E., Wheeler,E., Wong,S., Saudek,V. *et al.* (2014) Hypomorphism in human NSMCE2 linked to primordial dwarfism and insulin resistance. *J. Clin. Invest.*, **124**, 4028–4038.
65. Grange,L.J., Reynolds,J.J., Ullah,F., Isidor,B., Shearer,R.F., Latypova,X., Baxley,R.M., Oliver,A.W., Ganesh,A., Cooke,S.L. *et al.* (2022) Pathogenic variants in SLF2 and SMC5 cause segmented chromosomes and mosaic variegated hyperploidy. *Nat. Commun.*, **13**, 6664.
66. Pinto,A.E., Pereira,T., Santos,M., Branco,M., Dias,A., Silva,G.L., Ferreira,M.C. and André,S. (2013) DNA ploidy is an independent predictor of survival in breast invasive ductal carcinoma: a long-term multivariate analysis of 393 patients. *Ann. Surg. Oncol.*, **20**, 1530–1537.
67. Martínez-Jabaloyas,J.M., Ruiz-Cerdá,J.L., Hernández,M., Jiménez,A. and Jiménez-Cruz,F. (2002) Prognostic value of DNA ploidy and nuclear morphometry in prostate cancer treated with androgen deprivation. *Urology*, **59**, 715–720.
68. Zapatka,M., Pociño-Merino,I., Heluani-Gahete,H., Bermúdez-López,M., Tarrés,M., Ibars,E., Solé-Soler,R., Gutiérrez-Escribano,P., Apostolova,S., Casas,C. *et al.* (2019) Sumoylation of Smc5 promotes error-free bypass at damaged replication forks. *Cell Rep.*, **29**, 3160–3172.
69. Verver,D.E., Zheng,Y., Speijer,D., Hoebe,R., Dekker,H.L., Repping,S., Stap,J. and Hamer,G. (2016) Non-SMC element 2 (NSMCE2) of the SMC5/6 complex helps to resolve topological stress. *Int. J. Mol. Sci.*, **17**, 1782.
70. Ni,H.-J., Chang,Y.-N., Kao,P.-H., Chai,S.-P., Hsieh,Y.-H., Wang,D.-H. and Fong,J.C. (2012) Depletion of SUMO ligase hMMS21 impairs G1 to S transition in MCF-7 breast cancer cells. *Biochim. Biophys. Gen. Sub.*, **1820**, 1893–1900.
71. Graham,M.K. and Meeker,A. (2017) Telomeres and telomerase in prostate cancer development and therapy. *Nat. Rev. Urol.*, **14**, 607–619.
72. Storlazzi,C.T., Fioretos,T., Paulsson,K., Strömbeck,B., Lassen,C., Ahlgren,T., Juliusson,G., Mitelman,F., Rocchi,M. and Johansson,B. (2004) Identification of a commonly amplified 4.3 mb region with overexpression of C8FW, but not MYC in MYC-containing double minutes in myeloid malignancies. *Hum. Mol. Genet.*, **13**, 1479–1485.
73. Hanahan,D. and Weinberg,R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646–674.
74. Papp,B., Pál,C. and Hurst,L.D. (2003) Dosage sensitivity and the evolution of gene families in yeast. *Nature*, **424**, 194–197.
75. Zhang,S., Qi,Y., Liu,M. and Yang,C. (2013) SUMO E3 ligase AtMMS21 regulates drought tolerance in Arabidopsis thaliana(F). *J. Integr. Plant Biol.*, **55**, 83–95.
76. Lee,K.-J., Kim,Y.-E., Lee,H. and Park,S.-Y. (2017) Overexpression of SUMO E3 ligase HPY2 regulates the cell cycle in petunia development. *Hortic. Environ. Biotechnol.*, **58**, 384–392.
77. Spring,K., Ahangari,F., Scott,S.P., Waring,P., Purdie,D.M., Chen,P.C., Hourigan,K., Ramsay,J., McKinnon,P.J., Swift,M. *et al.* (2002) Mice heterozygous for mutation in Atm, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. *Nat. Genet.*, **32**, 185–190.
78. Goss,K.H., Risinger,M.A., Kordich,J.J., Sanz,M.M., Straughen,J.E., Slovek,L.E., Capobianco,A.J., German,J., Boivin,G.P. and Groden,J.

(2002) Enhanced tumor formation in mice heterozygous for Blm mutation. *Science*, **297**, 2051–2053.
79. Stopsack,K.H., Whittaker,C.A., Gerke,T.A., Loda,M., Kantoff,P.W., Mucci,L.A. and Amon,A. (2019) Aneuploidy drives lethal progression in prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 11390–11395.
80. Ippolito,M.R., Martis,V., Martin,S., Tijhuis,A.E., Hong,C., Wardenaar,R., Dumont,M., Zerbib,J., Spierings,D.C.J., Fachinetti,D. *et al.* (2021) Gene copy-number changes and chromosomal instability induced by aneuploidy confer resistance to chemotherapy. *Dev. Cell*, **56**, 2440–2454.
81. Lukow,D.A., Sausville,E.L., Suri,P., Chunduri,N.K., Wieland,A., Leu,J., Smith,J.C., Girish,V., Kumar,A.A., Kendall,J. *et al.* (2021) Chromosomal instability accelerates the evolution of resistance to anti-cancer therapies. *Dev. Cell*, **56**, 2427–2439.
82. Lafuente-Barquero,J., Luke-Glaser,S., Graf,M., Silva,S., Gómez-González,B., Lockhart,A., Lisby,M., Aguilera,A. and Luke,B. (2017) The Smc5/6 complex regulates the yeast Mph1 helicase at RNA-DNA hybrid-mediated DNA damage. *PLos Genet.*, **13**, e1007136.
83. Agashe,S., Joseph,C.R., Reyes,T.A.C., Menolfi,D., Giannattasio,M., Waizenegger,A., Szakal,B. and Branzei,D. (2021) Smc5/6 functions with Sgs1-Top3-Rmi1 to complete chromosome replication at natural pause sites. *Nat. Commun.*, **12**, 2111.
84. Madaan,K., Kaushik,D. and Verma,T. (2012) Hydroxyurea: a key player in cancer chemotherapy. *Expert Rev. Anticancer Ther.*, **12**, 19–29.
85. Saban,N. and Bujak,M. (2009) Hydroxyurea and hydroxamic acid derivatives as antitumor drugs. *Cancer Chemother. Pharmacol.*, **64**, 213–221.
86. Nazareth,D., Jones,M.J.K. and Gabrielli,B. (2019) Everything in moderation: lessons learned by exploiting moderate replication stress in cancer. *Cancers*, **11**, 1320.
87. Oo,Z.Y., Proctor,M., Stevenson,A.J., Nazareth,D., Fernando,M., Daignault,S.M., Lanagan,C., Walpole,S., Bonazzi,V., Škalamera,D. *et al.* (2019) Combined use of subclinical hydroxyurea and CHK1 inhibitor effectively controls melanoma and lung cancer progression, with reduced normal tissue toxicity compared to gemcitabine. *Mol. Oncol.*, **13**, 1503–1518.
88. Båvner,A., Matthews,J., Sanyal,S., Gustafsson,J.-A. and Treuter,E. (2005) EID3 is a novel EID family member and an inhibitor of CBP-dependent co-activation. *Nucleic Acids Res.*, **33**, 3561–3569.